#### DOCUMENT RESUME

ED 065 507

TM 001 427

AUTHOR

Engel, John D.

TITLE

Development of a Work Sample Criterion for General

Vehicle Mechanic.

INSTITUTION

Human Resources Research Organization, Fort Knox, Ky.

Div. 2.

SPONS AGENCY

Office of the Chief of Research and Development

(Army), Washington, D.C.

REPORT NO PUB DATE

TR-70-11 Jul 70

NOTE

39p.

EDRS PRICE

MF-\$0.65 HC-\$3.29

DESCRIPTORS

\*Auto Mechanics; \*Criterion Referenced Tests; Equivalency Tests; Item Analysis; \*Measurement Techniques; \*Performance Tests; Questionnaires;

Research Methodology; Skill Analysis; Task

Performance: \*Test Construction; Test Reliability;

Test Results

IDENTIFIERS

\*JOBTEST 1

#### ABSTRACT

A work sample criterion test was developed for General Vehicle Repairman, MOS 63C30 and 63C40. Test items covered three task categories: troubleshooting, corrective action, and preventive maintenance. Thirty-eight organizational mechanics were tested at Fort Knox, Kentucky. Data were also collected on the quality of performance, for example, use of good procedures, use of test equipment, and so forth. The study indicated that (a) the test appears to have a high degree of reliability (r=.82), (b) on the average, 60% of the test exercises were successfully completed by the 38 mechanics, (c) there was a moderate relationship between performance and length of experience, and (d) there were indications of lack of use and unfamiliarity with technical publications, and also a lack of skill in the use of special tools and equipment. (Author)

# **FD 065507**

**Technical Report 70-11** 

Development of a Work Sample Criterion for General Vehicle Mechanic

by

John D. Engel

**HumRRO** Division No. 2

AD\_\_\_\_\_

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM
THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY
REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

July 1970

Prepared for:

Office, Chief of Research and Development Department of the Army

Contract DAHC 19-70-C-0012

•

HUMAN RESOURCES RESEARCH ORGANIZATION

This document has been approved for public release and sale; its distribution is unlimited.

Destroy this report when it is no longer needed.

Do not return it to the originator.

### Development of a Work Sample Criterion for General Vehicle Mechanic

by

John D. Engel

This document has been approved for public release and sale; its distribution is unlimited.

**July 1970** 

Prepared for:
Office, Chief of Research and Development
Department of the Army
Contract DAHC 19-70-C-0012 (DA Proj 2Q062107A712)

HumRRO Division No. 2
Fort Knox, Kentucky
HUMAN RESOURCES RESEARCH ORGANIZATION

Technical Report 70-11 Work Unit JOBTEST Sub-Unit I



The Human Resources Research Organization (HumRRO) is a nonprofit corporation established in 1969 to conduct research in the field of training and education. It is a continuation of The George Washington University Human Resources Research Office. HumRRO's general purpose is to improve human performance, particularly in organizational settings, through behavioral and social science research, development, and consultation. HumRRO's mission in work performed under contract with the Department of the Army is to conduct research in the fields of training, motivation, and leadership.

The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

Published
July 1970
by
HUMAN RESOURCES RESEARCH ORGANIZATION
300 North Washington Street
Alexandria, Virginia 22314

Distributed under the authority of the Chief of Research and Development Department of the Army Washington, D.C. 20310



#### **FOREWORD**

The major long-term objective of Work Unit JOBTEST is to investigate and evaluate a variety of concepts and procedures for the measurement of job performances. Emphasis will be placed on identifying those techniques that have both validity and utility in practical testing environments and that have generality across groups of tasks.

JOBTEST I was concerned with the development of a "hands-on-equipment" work sample criterion in the area of automotive maintenance. This report describes the results of this work. In addition, information was gathered concerning the present level of performance of general vehicle repairmen, MOS 63C30 and 63C40. The research was performed and most of the report preparation completed while HumRRO was part of The George Washington University.

The research reported here is from the first of a series of studies. Subsequent experimental work will concern other measurement techniques in relation to the present work sample criterion test.

JOBTEST I was conducted during 1968 at HumRRO Division No. 2, Fort Knox, Kentucky, under Dr. Donald F. Haggard as Director. The Work Unit Leader is Mr. John

Military support for the Work Unit is provided by the U.S. Army Armor Center and by the U.S. Army Armor Human Research Unit; the Military Chief of the Unit is LTC John A. Hutchins. SP5 Robert Rehder of the Armor HRU served as research assistant in the study.

HumRRO research for the Department of the Army is conducted under Contract DAHC 19-70-C-0012. Work Unit JOBTEST is conducted under Army Project 2Q062107A712, Training, Motivation, and Leadership Research.

Meredith P. Crawford
President
Human Resources Research Organization



#### SUMMARY AND CONCLUSIONS

#### **MILITARY PROBLEM**

The Enlisted Personnel Management System is one of the largest Army users of proficiency tests. These tests are used to help implement many Army personnel programs, such as the proficiency pay and MOS qualification programs.

The Board of Inquiry on the Army Logistics System (Brown Board) and other surveys of maintenance activities have raised questions about the validity of proficiency measures used in the awarding of MOSs and proficiency pay for the automotive mechanic. It is considered likely that the method of proficiency measurement most commonly employed—paper-and-pencil examinations—limits the validity of proficiency assessment that may be achieved in evaluating certain types of tasks.

#### RESEARCH PROBLEM

The major long-term research problem in Work Unit JOBTEST is to study and evaluate a variety of concepts and procedures for the measurement of job performance. Emphasis will be placed on identifying those techniques that have both validity and utility in practical testing situations, and that have generality across groups of tasks.

The first phase of the research, and the primary problem dealt with in this report, is the development of a relevant and reliable work sample criterion for the General Vehicle Mechanic. This criterion will be used as a standard in later research phases that will evaluate various measurement techniques.

During the development of the criterion, a secondary objective was to collect information on the quality of performance of organizational maintenance by general vehicle repairmen.

#### **METHOD**

Work was begun by updating job information in a 1964 HumRRO analysis of job requirements for consolidated MOS 630, 631, 632 (Automotive Mechanic). This updated job requirements inventory was used as a basis for developing items for a "hands-on-equipment" work sample.

A four-day proficiency test consisting of 33 sample exercises was constructed. The test included a diagnostic scoring procedure for use in scoring men on quality of performance. The exercises were individually performed on track and wheel vehicles in common use and were individually scored by experienced mechanics who had been trained in proper test administration procedures.

The test was administered to a total of 38 organizational mechanics, drawn from all the organizational maintenance units at Fort Knox, Kentucky. In addition, a question-naire was used to obtain information on personnel data, organizational maintenance experience, experience on various vehicle systems, current job assignment, type and amount of training, and amount of supervision received on the job.

#### **RESULTS**

(1) The item analysis values for the test were found to be well within accepted ranges for this type of data analysis. These analyses dealt with the difficulty level of the



items and their ability to discriminate between subjects who were high and low on the total pool of items.

(2) The total test appears to have a high degree of reliability (r=.82), indicating it should permit a high degree of accuracy of measurement when used as a criterion in evaluating other measurement techniques.

(3) On the average, 60% of the exercises were successfully completed by the 38 mechanics.

(4) There was a moderate relationship between performance and length of experience.

(5) There was no practical difference in performance among mechanics with different types of training (e.g., service school, other school, on-the-job).

(6) Lack of use and unfamiliarity with technical publications, and lack of skill in the use of special tools and equipment were significant correlates of low proficiency.

#### **CONCLUSIONS**

- (1) The work sample criterion appears to be a reliable and job relevant measure that may be used as a standard for the evaluation of other measurement techniques.
- (2) The results of this study indicate that the development of a proficient, well-trained mechanic is extremely complex, and that a more detailed examination should be made in the areas of training for troubleshooting tasks (as these appear to be the most difficult tasks to perform) and training in the use of technical publications and test equipment.



vi

#### CONTENTS

		age
Introdu		3
	Military Problem	3
	An Overview of Proficiency Measurement Techniques	3
	Research Problem	5
	Approach	6
Method		6
	Subjects	6
	Equipment, Materials, and Research Personnel	7
	Test Development	7
	Test Item Format	8
	Test Administrator Training	9
	Test Administration	46
	Test Scores and Personnel Data	10
Results		10
	Test Characteristics	10
	Item Analysis	10
	Test Reliability	12
	Performance of Organizational Maintenance by General Vehicle Repairmen	13
Discuss	sion and Observations	16
1.20	Oland	21
Literatu	re Cited	21
Append	ices	
A	Proficiency Test for General Vehicle Repairman: List of Exercises	23
В	Proficiency Test for General Vehicle Repairman: Sample Exercise	
С	Proficiency Test for General Vehicle Repairman: Sample Summary	
	Performance Checklist	26
D	Proficiency Test for General Vehicle Repairman: Sample Organizational	
	Maintenance Personnel Inventory	27
Ε	Difficulty Level (P) and Discrimination Index $(r_b)$ of Test Items for the	
_	Total Test and Subtests	29
		_
Tables		
1	Number of Subjects Tested in Each MOS by Type of Training and Years of Experience	(
2	Range and Mean Difficulty Level (P) Values for Various Groups of Test Items	11
3	Range and Mean of the Discrimination Indices Between Test Items and	
	Various Test Scores	1
4	Average Difficulty Level (P) and Mean Scores for Half Tests	12
5	Mean Test Performance by MOS and Total Group	13
6	Correlation Between Work Sample Test Scores and Length of Experience	14
7	Mean Test Performance by Various Types of Training	
•	The second secon	٠



vii

Tables		Page
8	Tests of Differences Between the GT and MM Means for Upper and	
	Lower Performance Groups	15
9	Test Item Performance in the Descriptive Pass-Fail Categories	15
10	Percent of Test Items in Which "Required" Procedures Were Used by	
	High and Low Performance Groups	15
11	Percent of Test Items in Categories of Test Equipment Usage for High and	
	Low Performance Groups	16
12	Percent of Test Items in Categories of Publications Usage for High and	
	Low Performance Groups	16



Development of a Work Sample Criterion for General Vehicle Mechanic



#### INTRODUCTION

#### **MILITARY PROBLEM**

A primary reason for the use of proficiency tests in the U.S. Army is to support the implementation of certain portions of the Enlisted Personnel Management System. The Management System, in turn, was devised to implement certain portions of the mission of the Deputy Chief of Staff for Personnel, Department of the Army (DCSPER), specifically DCSPER's responsibility for "development and administration of a military personnel management system, to include policies and programs for procurement, individual training (less foreign military training), education, retention, career development, distribution, promotion, and separation of military personnel . . . "(1) The present study is based on a requirement from DCSPER.

One of the largest users of proficiency tests is the Enlisted Personnel Management System. A survey by the Board of Inquiry on the Army Logistics System (Brown Board) and other surveys of maintenance activities have raised questions about the validity of the proficiency measures and performance standards employed in the Enlisted Evaluation System's MOS evaluation.

These MOS evaluation results are used in numerous Army personnel programs, some examples of which are:

- (1) Proficiency pay programs: MOS evaluation is designed to provide incentives to keep qualified soldiers with critical occupational skills.
- (2) MOS qualification: MOS evaluation provides for the verification of each soldier's job proficiency.
- (3) Secondary MOS qualification: MOS evaluation ensures that the soldier has maintained his job proficiency depth in effective assignment and utilization actions.
- (4) Reserve and National Guard: Annual MOS evaluations of Reserve Component Unit Personnel help ascertain the readiness posture of our civilian soldiers.
- (5) Promotion qualification: MOS evaluation is geared to determine the enlisted man's promotion qualification. The use of the promotion qualification score as a criterion for promotion is mandatory.
- (6) Enlisted grade and MOS determination: MOS evaluation is used to determine enlisted grade and MOS for those officers and warrant officers who are released from active duty in their commissioned or warrant officer status and enlist in the Regular Army.

In general, the Enlisted Personnel Management's Enlisted Evaluation System needs job proficiency tests that have a high degree of validity and reliability, and that are as inexpensive to develop and administer as possible (consonant with standards of validity and reliability), in order to effectively implement several portions of the Army's personnel program.

#### AN OVERVIEW OF PROFICIENCY MEASUREMENT TECHNIQUES

There are various types of proficiency measurement techniques that might be used to evaluate a man's achievement. According to Glaser and Klaus (2), proficiency measurement techniques may be grossly categorized on the basis of their remoteness from actual



job performance. This remoteness may be due to differences in (a) the behavior elicited for measurement, (b) the eliciting stimuli themselves, or (c) both stimulus and behavior. In most instances, however, as the test stimuli become more remote from those found in the actual job situation, the responses elicited are likewise less similar to those found in job performance.

Thus at one extreme along this continuum of remoteness is the measurement of proficiency during actual job performance. At the other extreme are measures (e.g., paper-and-pencil tests) that are not obviously similar to the criterion task, but assess performance at tasks that correlate with on-the-job behavior.

Between these two extremes are test situations that (a) call for the performance of the actual job task outside the real job environment, or (b) attempt to simulate the job task while at the same time offering effective control of the factors that in "real" situations are likely to interfere with reliable and valid measurement. The four major segments along this continuum can be identified as (a) on-the-job measures, (b) work sample measures, (c) simulated-job measures, and (d) correlated-job measures.

In principle, proficiency measurement should be accomplished during a man's typical performance, under conditions generally present during day-to-day operations. This method, however, presents a number of problems. The degree of control that can be achieved in a job situation is generally less than satisfactory for obtaining reliable measurements. In addition, attempts to standardize the situation for proficiency-measurement purposes frequently introduce considerable artificiality into the situation. Finally, the consideration of committing large amounts of time, money, and men to the testing situation often makes this an impractical method of assessment.

To reduce, to some extent, the problems involved in on-the-job measures, samples of the actual job tasks involved may be removed from the real job environment so they can be readily and reliably assessed. This type of proficiency measurement technique is referred to as a work sample test. Here, the individual performs the actual tasks but not in the real job environment. This technique is a close approximation to on-the-job measures, but it has some of the same drawbacks: It is costly, time consuming, and essentially impractical as a method of assessing large numbers of people.

Because it is difficult to measure men's proficiency during actual job and work sample situations, the job must be simulated in a controlled manner in order to produce a reliable and valid, yet practical method of performance assessment. The essence of task simulation is the design of test stimuli that will evoke joblike responses that can be measured objectively. This general category of simulated-job measures includes a variety of proficiency measurement techniques. Some of the most frequently employed measures use equipment mock-ups and simulators.

An extreme position along the dimension of remoteness from job reality is represented by tests measuring, not job behaviors themselves, but correlated-job behaviors—i.e., measures correlated with job behavior. These measures are the most remote from the actual job situation. The most widespread type of correlated-job measure is verbal response as used to assess skills that are substantially nonverbal. Examples of this type of proficiency measure are tests of job knowledge, vocabulary, and nomenclature used to evaluate performance at procedural and manipulative tasks. Other types of correlated-job measures are those that involve a deliberate modification in the response made so as to facilitate the recording and evaluation of responses. A common example of this kind of construction is multiple-choice paper-and-pencil tests that are used to measure the ability to produce appropriate responses by measuring ability to recognize them. Because they are easily constructed, inexpensive, and easily administered, paper-and-pencil tests of job knowledge are frequently used to evaluate an individual's proficiency.

However, tests measuring knowledge of technical information, tool nomenclature, technical vocabulary, or underlying theory may not relate to actual performance for some

ERIC Full Text Provided by ERIC

tasks. Instead, they measure verbal knowledge about the job, and therefore assess behaviors which, at best, may be correlated only slightly with actual job behaviorespecially if the job depends on motor and manipulative skills.

It is thus likely that paper-and-pencil tests, the kind of proficiency measurement most commonly employed, limit the validity of the evaluation of proficiency for certain tasks. Two examples are summarized below.

(1) The U.S. Army Enlisted Evaluation Center conducted a validation study of the paper-and-pencil evaluation test for Track Vehicle Mechanic, MOS 63C20 (Yellen, 3). Performance of duties in this MOS is heavily dependent on perceptual-motor skills. The criterion used in the study was the average co-worker rating rendered by three enlisted men for each of the 47 enlisted men in the validation sample. The following conclusions were drawn from the study: "(a) The total evaluation test had a validity coefficient of .21; (b) optimal weighting of the Broad Subject Matter Areas did not significantly increase the validity of the total evaluation test."

It should be noted that a validity coefficient r of .21 is not statistically significant, but even if it were, it is far too low for use in group measurement and grossly inadequate for use in individual measurement.

One reason for the lack of validity could be that an examination can differentially affect scores because of factors unrelated to actual job performance. The test may lean heavily, for example, on the individual's ability to understand test directions or on his speed in reading lengthy descriptive passages. The influence of verbal facility on the test score appears to distort proficiency estimates in such a way as to systematically penalize those with poor vocabularies and reading skills, and not to reflect their proficiency in tasks that do not depend directly on verbal skills.

(2) The Enlisted Evaluation Center conducted a validation study of the paperand-pencil evaluation test for Personnel Specialist, MOS 716.1 (Urry, Shirkey, and Nicewander, 4). In this MOS, performance of duties is heavily dependent on verbal and reasoning skills, in definite contrast to the previous case in which perceptual-motor skills were predominant. Again, the criterion used in the study was the average co-worker rating rendered by three enlisted men for each of the 55 enlisted men included in the validation sample. The following conclusions were drawn from the study: "(a) The total evaluation test had substantial validity—the validity coefficient was .50; (b) optimal weighting of BSMAs did not increase the validity of the total evaluation test."

It should be noted that a validity coefficient r of .50 is sufficiently high for practical application in group measurement.

One reason for the validity of the test for Personnel Specialists may be the fact that the test leans heavily upon the same skills that are required in the job-both the test and the job rely directly on verbal and reasoning skills. Therefore, in this instance a paper-and-pencil test seems to be appropriate for measuring performance in the MOS.

#### RESEARCH PROBLEM

The long-range research problem of Work Unit JOBTEST is to study and evaluate a variety of concepts and procedures for the measurement of job performance. Emphasis will be placed on identifying those techniques which have both validity and utility in practical testing environments, and which have generality across groups of tasks.

The primary research problem dealt with in this report was to develop a relevant and reliable work sample criterion, which may be used as a standard against which all techniques will be evaluated. A secondary aspect deals with obtaining information about the quality of performance of organizational maintenance by general vehicle repairmen.



#### **APPROACH**

A two-step approach was taken in order to achieve the long-range research objective. This report deals with the first step or the immediate research objective, that is, the development of a reliable work sample for assessing on-job performance, since performance on the job must serve as the criterion for the validation of other assessment techniques.

The second step will be concerned with the research and development of a range of proficiency measurement techniques varying in degree of remoteness from the work sample. These techniques will be compared with the work sample standard (as described in this report) on the basis of reliability, validity, ease of manufacture, application, and cost.

#### METHOD

#### **SUBJECTS**

The G3 Personnel Section at Fort Knox, Kentucky, assigned 38 organizational mechanics from the available manpower in the various organizational maintenance units there. The selection was further based on MOS Code, and every effort was made by G3 to obtain men with varying degrees of experience and training within the 63C30 and 63C40 MOS Codes. Fifteen men held the MOS Code 63C30 and 23 the Code 63C40 (see Table 1). The subjects were divided into four groups—two of 63C30s and two of 63C40s, each group being tested for four days.

Table 1

Number of Subjects Tested in Each MOS by Type of Training and Years of Experience

Type of Training	Experience (years)					
and MOS	0-1	1-5	5-10	10-15	15 or More	Total
Service School						
MOS 63C30	0	0	2	0	0	2
MOS 63C40	0	3	2	2	0	7
Other School						
MOS 63C30	1	3	1	0	0	5
MOS 63C40	0	2	2	3	3	10
Service and						
Other School						
MOS 63C30	1	5	0	1	0	7
MOS 63C40	1	0	0	1	0	2
On-the-Job						
MOS 63C30	. 0	1	0	0	0	1
MOS 63C40	0	2	1	0	1	4
Total	3	16	8	7	4	38



#### EQUIPMENT, MATERIALS, AND RESEARCH PERSONNEL

Test problems were set into the following vehicles:

Quantit	y T	ype
3	M60	tanks
6	M151A	l ¼-ton trucks
2	M54A2	5-ton trucks
1	M35A1	2½-ton truck
1	M88	recovery vehicle
1	M110	self-propelled artillery vehicle
1	M113	armored personnel carrier
3	M108	self-propelled artillery vehicles

Total 18

Two large maintenance shops, each with two indoor and three outdoor bays, were used as testing facilities.

Each man was provided with a mechanic's tool kit, relevant technical manuals, troubleshooting guides, and test equipment. Test item sheets, scoring sheets, and a biographical questionnaire were developed (copies of these items are shown in Appendices A through D).

Ten mechanics, E5 or higher, with the MOS Code 63C30 or 63C40, were trained to be test administrators. In addition, one noncommissioned officer in charge (NCOIC), one warrant officer, and one civilian researcher formed the research team and were responsible for the testing.

#### **TEST DEVELOPMENT**

A task inventory was necessary, to serve as a basis for future comparisons and as a reference for the development and evaluation of new performance measures. Such a task inventory had been completed at HumRRO Division No. 2 in March 1964, and it was considered sufficiently detailed and inclusive to warrant updating and use for the present study.<sup>1</sup>

Three senior automotive mechanics, with ten or more years of experience, served as technical advisors to review and update the 1964 inventory, with the constraints that (a) only task categories pertaining to individual (as opposed to group) performance were to be considered; and (b) only job level "2" tasks (defined in the task inventory as essentially journeyman level tasks) were to be included.

These constraints delimit the study and focus on tasks to be undertaken by journeymen—the level at which the mechanic is first responsible for performance of the full spectrum of organizational maintenance tasks.

In the review and alteration of the Smith inventory, each of the three technical advisors was given a copy of the complete inventory and asked to pencil in any revisions deemed necessary in his independent appraisal. The reports were then combined for comparison, and the technical advisors' modal responses were annotated for each area.

On the basis of the information gathered (5), activities to be tested were selected from three task categories: (a) troubleshooting, (b) corrective action, and (c) preventive maintenance. In addition, examination of the task inventory showed that approximately

ERIC \*\*
Full Text Provided by ERIC

15

<sup>&</sup>lt;sup>1</sup> The analysis for the task inventory was performed by Dr. John P. Smith.

60% of the vehicle systems on which the *individual* journeyman mechanic worked contained troubleshooting tasks; 35% of the vehicle systems on which the *individual* worked contained corrective action tasks; and 5% of the vehicle systems on which the *individual* worked contained preventive maintenance tasks.

Test problems were drafted by one automotive Warrant Officer and two E7 mechanics. Each man drafted 11 test items that were based on tasks he chose because of his expertise in specific automotive areas. These test items were drafted using the information provided in the job requirements inventory and the following criteria:

- (1) All items were to cover *individual* (as opposed to group) task categories of the journeyman level mechanic.
- (2) Problems covering the various task categories were to be represented in approximately the same proportion as the task categories occur in the job requirements inventory.
- (3) Approximately half of the items were to deal with wheel vehicle maintenance, the other half with track vehicle maintenance.
- (4) The task which the item was based on should be critical. That is, (a) without proper maintenance of part or system, vehicle could not maintain "combat readiness" (i.e., the vehicle must be able to make a trip of approximately 85 miles in not more than 15 hours); (b) if malfunction were not properly identified and diagnosed, a costly part or system of the vehicle might be unnecessarily replaced. Only tasks designated as "High" on criticality would be considered.
- (5) The problem should be one that often occurs in the field, that is, a problem of high or medium frequency. Frequency was defined as:

Low: No more than once or twice a year in a battalion-size unit.

High: More than six times a year in a battalion-size unit.

(6) High-density vehicles or vehicle systems would be utilized in forming problems.

These criteria were carefully considered in preparing each item, but not all of them were maximally met in each problem. The various criteria had to be weighed by the item writers in order to produce a representative sample of items which could be used to test as many people as possible within a given period of time and within the limits imposed by availability of equipment and personnel.

A total of 33 items—17 troubleshooting, 12 corrective action, and four preventive maintenance— were completed and sent to the Automotive Department of the U.S. Army Armor School to be reviewed for currency and technical accuracy.

#### **TEST ITEM FORMAT**

In order to standardize test administration and scoring procedures, each item was composed of two parts. The first part (Appendix B) presented the content. It began with a statement of the symptom (for troubleshooting problems) or the action to be performed (for corrective action and preventive maintenance problems); this section was read by the tester to the subject. The essential procedures or approved steps for completing the task were then listed, with each step keyed as either essential (E) or optional (O), according to the judgment of the technical experts. Also included were spaces for writing in any parts mistakenly identified for replacement, and for recording the subject's previous experience with similar tasks.

The second part (Appendix C) was a performance checklist common to all test items. It was headed by a block of four passing and four failing categories. If a subject correctly reported the malfunction or correctly performed the required action, he was

ERIC

Full Text Provided by ERIC

scored in the pass section, and his degree of pass was indicated by one of four descriptive statements. The same type of scoring was used if he failed the problem. These descriptive statements are:

#### Pass:

Used good procedures; knew what he was doing Had a pretty good idea, but some guess work Knew very little; probably just a good guess Knew very little; but used publication accurately

#### Fail:

Didn't know enough to get started Started but gave up quickly Some knowledge of what to do but much guess work also Had fairly good grasp of the problem but failed

The use of this scoring provided additional information on the quality of performance which could be related to the Army's training system, in terms of identifying areas for training emphasis.

All of the test items were scored on this two-part form. In this way, testing and scoring were made fairly uniform.

#### **TEST ADMINISTRATOR TRAINING**

Ten experienced mechanics (E5s and E6s) were trained to administer the tests. The training period for the testers was three and one-half days, during which time vehicles were placed in the working bays and problems were set into them. The problems were then tried out by the mechanics so they could familiarize themselves with test administration procedures and typical problems that might be encountered.

#### **TEST ADMINISTRATION**

Before a subject was tested, he was asked to fill out a questionnaire (shown in Appendix D) that required information in the following areas: (a) personnel data, (b) organizational maintenance experience on track and wheel vehicles, (c) experience on various vehicle systems, (d) current job assignment, (e) type and amount of training, and (f) amount of supervision received on the job.

The subjects then received a thorough briefing on the nature of the tests and test procedures, and were encouraged to do their best on each problem. It was emphasized that they were not being personally evaluated for the record, but that their scores would be used solely for research purposes. They were further encouraged to view the tests as a novel learning experience.

During the first day and a half of testing, each man was given the 17 trouble-shooting problems (20 minutes per station). During the next day and a half, they were given the 12 corrective action problems (20 minutes per station for eight stations, and 30 minutes per station for four stations) and during the last day, the four preventive maintenance problems (15 minutes per station).

Each day, the individual was given a schedule of his route from station to station and was told which station to report to first. One vehicle was located in each of five bays



. 17

(or work areas) of two large motor-pool buildings.<sup>2</sup> At each station the pretrained tester read the specific directions for the problem to the subject. (The directions were printed on each test sheet.) The subject was further instructed to proceed as he normally would on the job, using the test equipment, tools, and publications located at each test station. Testers were instructed to avoid conversation with the men being examined and to give them no help beyond repeating instructions. However, the tester did act as a helper by cranking the engine, holding a light, turning switches, and so forth, when requested.

Performance was observed by the tester and recorded on the checklist. At the end of the time period (20 or 30 minutes), an air horn was blown and the men moved to another station, in this way completing the prescribed number of tests. After each exercise, the tester completed the summary score sheet (shown in Appendix C) which contained pass-fail categories of performance and some 20 statements describing the errors of performance. This same overall testing procedure was used for all four groups.

The test stations were situated in such a way that a man could not easily observe the activity at another test station. In addition, men in the four different test groups were selected, whenever possible, from different organizational maintenance units at Fort Knox, so that communication between groups was minimized.

It is believed that the actual job tasks were meaningfully duplicated except in two cases where it was necessary to remove the power plants from the vehicles and hook them up to the power source with "slave" cables or "ground-hop" kits.

#### **TEST SCORES AND PERSONNEL DATA**

Each exercise was scored on a pass or fail basis. A man's total score was simply the number of exercises he completed correctly out of the 33 on which he was tested.

The supplementary data used in the analysis include: (a) General Technical (GT) and Motor Maintenance (MM) aptitude scores;<sup>3</sup> (b) number of years of maintenance experience; (c) type and number of maintenance schools completed; (d) current job assignment; (e) experience on different types of vehicles; and (f) supervision given on the job.

#### **RESULTS**

#### **TEST CHARACTERISTICS**

#### Item Analysis

While a preliminary item analysis was performed, it should be realized that item analysis for this criterion measure is not the same as for classical psychometric applications, because the content of a job sample criterion is established by analysis of the job. In other applications of item analysis, for classical psychometric uses, the main purpose is establishing a homogeneous pool of items. Homogeneity in the pool of items is, in principle, inapplicable for work sample tests such as this criterion measure whose content is job sampling or measuring components of a job which may or may not form a



<sup>&</sup>lt;sup>2</sup>The other vehicles were parked in a third motor-pool building and used as a back-up for vehicles which may have become defective during testing.

<sup>&</sup>lt;sup>3</sup>Army Aptitude Area scores based on combinations of scores on the Army Classification Battery (ACB).

homogeneous pool of items. Consequently, the item analysis used in this study provides data on the nature of the items rather than a direct and specific basis for eliminating items from a final form of the test.<sup>4</sup>

The item analysis was concerned with two statistical aspects: first, the difficulty level of the item, that is, the proportion of people who get the item correct; second, the discrimination index of the item, that is, the degree to which the item differentiates subjects who are high from those who are low, in terms of a standard. The standard, in this case, was performance on the complete pool of items.

The range and mean values for the difficulty level (P) of the various groups of items are shown in Table 2, and a more complete table of all item P values is presented in Appendix E. These data indicate that the troubleshooting items tend to be the most difficult, and the corrective action items the least difficult.

The second item analysis characteristic is the degree to which the item differentiates subjects who are high from those who are low on the complete pool of items. This index gives an indication of the internal consistency of the test. The range and mean discrimination indices of the items are given in Table 3, and a more complete table of all item discrimination indices is presented in

Table 2

Range and Mean Difficulty Level (P) Values for Various Groups of Test Items

T I	Difficulty Level (P)			
Test Items	Range	Mean		
Troubleshooting	.2489	.48		
Corrective Action	.32-1.00a	.77		
Preventive Maintenance	.3781	.63		
Total Test	.24-1.00b	.60		

<sup>a</sup>If the one item with a P value of 1.00 were eliminated, the range would be .32-.89.

bIf the one item with a P value of 1.00 were eliminated, the range would be .24-.89.

Appendix E. Table 3 and Appendix E include discrimination indices both in relation to total test score and in relation to the relevant subtest (e.g., troubleshooting) score. These data indicate that the corrective action items tend to show the greatest internal consistency, and the troubleshooting items the least consistency.

Table 3

Range and Mean of the Discrimination Indicesa

Between Test Items and Various Test Scores

				Tes	t Items	-			
Test Scores	Troubleshooting Co		Corrective	Corrective Action		Preventive Maintenance		All Test Items	
	Range	Mean	Range	Mean	Range	Mean	Range	Mean	
Subtest	1075	.30	.1988	.59	. 19 64	.36	-	_	
Total Test	.1083	.37	.1296	.65	.3653	.42	.1096	.48	

<sup>&</sup>lt;sup>a</sup>The Biserial Coefficient of Correlation has been corrected for spurious item-total overlap.

The lower item difficulty and discrimination indices for the troubleshooting items are most probably due to the greater heterogeneity in item content, and also in item difficulty, as compared to the same properties in the corrective action and preventive maintenance items.

Therefore, the results of the item analysis, as shown in Tables 2 and 3, indicate that: (a) The average difficulty level of the items in the test is acceptable, as is the range



<sup>&</sup>lt;sup>4</sup>These data would be useful as a means of discovering and eliminating unsatisfactory items rather than as a means of selecting a small fraction of items that can be identified as "the best."

of the difficulty indices (for efficient differential measurement, a test must have a relatively wide range of item difficulties with an average difficulty level of about .50). (b) The items are useful in distinguishing between those who score high and those who score low on the total test.

#### **Test Reliability**

It was necessary to establish the level of reliability of the present criterion, because a necessary (but not sufficient) condition for correlation between a criterion measure, such as the present work sample, and the theoretically perfect ultimate criterion of job success is that the criterion measure have some reliability (i.e., consistency with which the measuring instrument would produce the same measurement in testing and retesting an individual).

To extract an estimate of reliability from a single administration of the test, it was decided to divide the total test into two half-length tests and correlate the scores on the halves. The test was therefore divided so that the items in each half would be as comparable as possible on the following factors: (a) problem type (troubleshooting, corrective action, preventive maintenance); (b) vehicle type (wheel or track); (c) vehicle system (engine, cooling, electrical, etc.); and (d) special tools and equipment required to solve the problem.

Of the 12 items in the corrective action part of the test, one was eliminated because, after repeated work on the vehicle, the part to be repaired became clearly identifiable and therefore cued the subject to the correct action. Of the 11 items remaining, one had to be eliminated in order to make an equal distribution of items into two separate half tests. It was decided to eliminate the only item that showed an  $r_b$  greater than 1.00 because of unusual distribution features.

Table 4

Average Difficulty Level (P) and Mean Scores for Half Tests

	Test Characteristics			
Test Elements		Items Correct		
	Average P	Mean	Standard Deviation	
Troubleshooting Subtest				
Half-Test 1	.46	3.6	1.5	
Half-Test 2	.46	3.6	1.7	
Corrective Action Subtest				
Half-Test 1	.73	3.6	1.7	
Half-Test 2	.74	3.7	1.4	
Preventive Maintenance Subtest				
Half-Test 1	.67	1.0	0.7	
Half-Test 2	.59	1.5	0.8	
Total Test				
Half-Test 1	.57	8.6	3.1	
Half-Test 2	.57	8.6	2.9	

12

Of the 17 items in the troubleshooting part of the test, one had to be eliminated in order to make an equal distribution of the items into two separate half tests. It was decided to eliminate the item that had the lowest  $r_b$  and at the same time the most extreme P value.

After items were discarded and the remainder divided, each half test consisted of 15 items. The two half tests were then analyzed; the average difficulty level (P value) and mean items correct for each half test are given in Table 4.

The data indicate that the two half tests are not only theoretically equivalent but also statistically equivalent—that is, tests of significance of differences showed no significant difference between the means and standard deviations.

The scores on the two half tests were then correlated, yielding values (corrected for double-length) between the troubleshooting, corrective action, and preventive maintenance tests, respectively, of .71, .78 and .63. The correlation between the total score on both half tests was .82 (corrected for double-length). All these correlation coefficients indicate, both statistically and practically, a significantly high degree of test reliability.

These results therefore indicate that the work sample criterion appears to satisfy the requirement for "some degree of reliability."

#### PERFORMANCE OF ORGANIZATIONAL MAINTENANCE BY GENERAL VEHICLE REPAIRMEN

A secondary objective of the study was to gather information on the current performance level and quality of performance of automotive mechanics. This information is presented in the following sections.

Test Performance. The percent of total items passed by the group, by MOS skill level, is presented in Table 5.

There is no practical difference in performance between the higher and lower skill level mechanic as evidenced by the fact that 63C30 mechanics passed 60% of the items, and the 63C40 mechanics 61%. According to the existing MOS structure and philosophy, one might expect this result; the basic difference between an MOS 63C30 and an MOS 63C40 mechanic is in supervisory skills, which would not make a difference in the performance of hands-on-equipment test problems.

Table 5

Mean Test Performance by

MOS and Total Group

MOS	Number of Subjects	Mean Percent of Total Items Passed	
63C30	15	60	
63C40	23	61	
Total	38	60	

Maintenance Experience. When test performance was examined as a function of automotive maintenance experience, the data presented in Table 6 indicate only a moderate relationship between a man's length of experience on the job and his level of performance in the areas of troubleshooting and corrective action tasks. This relationship does not seem to hold for preventive maintenance tasks; however, the nature of such tasks would not require that a man have a great deal of experience to perform them adequately, so the lack of correlation is not surprising.

Thus, the data indicate that on the basis of a broad correlational analysis, there is a moderate relationship between length of experience and performance on the test.

Maintenance Training. Test performance was studied as a function of type of training (see Table 7). Mechanics who had some kind of formal training performed slightly better than mechanics who had only on-the-job training. On the basis of these data, there would seem to be little, if any, difference in performance among mechanics who have different types of formal training.



Table 6

Correlation Between Work Sample
Test Scores and Length of Experience
(N=38)

Test Scores	Correlation a With Experience	Р
Troubleshooting	.41	<.05
Corrective Action	.39	<.05
Preventive Maintenance	.04	NS
Total Test	.38	<.05

<sup>&</sup>lt;sup>a</sup>Pearson Product Moment Correlations were computed.

Job Assignment. The 34 subjects<sup>5</sup> were divided by current job assignment into two groups—maintenance job in primary MOS, and supervisory job in primary MOS. For the 15 men who were at that time assigned to a maintenance job in their primary MOS, the average performance score was 64%. For the 19 men assigned to a supervisory job in their primary MOS, the average performance score was 61%. Thus, there was little difference in the level of performance between people in different job assignments.

Aptitude Area Scores. The means of the subjects' General Technical (GT) and Motor Maintenance (MM) aptitude

scores were 97.9 and 110.4 respectively. Pearson Product Moment correlations were calculated between the aptitude scores and the performance test scores. The r between GT and performance test scores was .35 and the r between MM and work sample scores was .38. These correlations show a moderate degree of relationship between the aptitude variables and performance test proficiency; they are statistically significant (p < .05).

Table 7

Mean Test Performance by Various Types of Training

Type of Training and School	Mean Percent of Total Items Passed
Service School (Fort Knox, Fort Sill, Fort Benning)	66
Other School (USAREUR, Aberdeen Training Center)	63
Service and Other School (Combination of above schools)	55
On-the-Job-Training	49

The results of significance of difference tests between the GT and MM mean scores and for subjects who scored in the upper or in the lower third of the total group are provided in Table 8.

The results indicate that while the amount of difference was large enough to produce a significant difference between the high and low performance groups in mean GT scores, there was not sufficient difference to demonstrate difference in MM scores between the high and low performance groups.

Performance Characteristics. Performance characteristics of the upper and lower thirds of the group as described in the pass-fail categories are given in Table 9.

As the data indicate, a high percentage of those who failed were described by the statement: "Didn't know enough to get started." In contrast, to a high percentage of those who passed the statement, "Used good procedures, knew what he was doing," was applicable. It should be noted that among those people who failed and were in the upper one-third of the group, the greatest percentage still used "fairly good procedures."

<sup>&</sup>lt;sup>5</sup>Of the total 38 subjects tested, four were holding clerical jobs in their primary MOS and were therefore eliminated from this analysis.

Table 8

Test of Differences Between the GT and MM Means for Upper and Lower Performance Groups
(N=26)

Aptitude Test	Score Group	Mean	Standard Deviation	df	ι	p
GT	Upper Third Lower Third	100.1 89.6	2.6 4.2	24	2.12	<.05
MM	Upper Third Lower Third	114.4 105.6	3.7 4.8	24	1.44	NS

Use of Required Procedures. The degree to which "required" steps were followed in completing each test item is presented in Table 10. It would be expected that those who scored higher on the test also completed the "required" steps, and the data show this trend; however, the trend is not as strong as might be expected. If these steps, taught in the schools and listed in the technical publications, are truly "required," one would expect successful performance to be absolutely dependent on the completion of all the "required" steps. However, the data indicate that frequently men pass test items and perform less than 100% of the "required" procedures. Therefore, it may be concluded that these required steps are helpful but not essential to successful performance.

Table 9

Test Item Performance in the Descriptive Pass-Fail Categories

Category	Total Test Performance (percent)			
<b>.</b>	Upper Third	Lower Third		
Pass				
Good Procedures	63	23		
Some Guesswork	10	4		
All Guesswork	0	9		
Accurately Used	4	7		
Publications				
Fail				
Fairly Good Procedures	13	3		
Much Guesswork	5	8		
Gave up Quickly	1	2		
Didn't Get Started	4	43		

Use of Test Equipment. The use of test equipment by people in the upper and lower performance groups is shown in Table 11. Not all problems required the use of test

Table 10

Percent of Test Items in Which "Required" Procedures Were

Used by High and Low Performance Groups

m . ) m .	"Required" Procedures Followed							
Total Test Performance	More Than One-Half	One-Half	Less Than One-Half	None				
Upper Third	54	4	14	28				
Lower Third	34	3	11	52				

Table 11

Percent of Test Items in Categories of Test Equipment
Usage for High and Low Performance Groups

Total Test	Use of Test Equipment							
Performance	Used Properly	Did Not Use	Could Not Operate	Did Not Know Procedures				
Upper Third	54	41	1	3				
Lower Third	24	70	0	5				

equipment, and those that did not are combined, in the first category of the table, with the occurrence of proper usage when required.

As expected, the people in the high performance group used the test equipment far more often than those in the low performance group. Failure to use the test equipment may have been due either to not knowing how to use it or not knowing when to use it. The low performing group was primarily characterized by the statement, "did not use the equipment."

Use of Publications. The difference with which technical publications were used by the people who scored in the upper and lower thirds of the total group is shown in Table 12.

Table 12

Percent of Test Items in Categories of Publications Usage for High and Low Performance Groups

m 1 m	Use of Publications								
Total Test Performance	Used Properly	Did Not Use	Could Not Find Reference	Could Not Understand Reference					
Upper Third	53	21	14	12					
Lower Third	18	57	9	15					

Failure to use publications and their misuse occurred often in both groups, but far more often in the low performance group. Even though the high performance group used the manuals correctly on 53% of the items, there was still a surprisingly high 47% of the items in which the manuals were not used or were used incorrectly.

By contrast, the low performance group used the manuals correctly on only 18% of the items; on the rest of the items, the manuals either were not used or were used incorrectly.

#### DISCUSSION AND OBSERVATIONS

The results of this study indicate that the work sample criterion appears to be, in fact, a reliable (r=.82) and relevant measure of a mechanic's proficiency level. This work sample may consequently be used as a standard for the evaluation of other measurement techniques.

Concerning the performance of organizational maintenance by general vehicle repairmen, in many respects the results of the current testing program tend to corroborate the findings of similar studies on organizational mechanics. In 1966,<sup>6</sup> the performance level of organizational mechanics was found to be approximately 68%; the current study found their performance level to be approximately 60%. Lack of use and unfamiliarity with technical publications, in addition to lack of skill in the use of special tools and equipment, continue to be significant correlates of low proficiency. There is also a moderate correlation between experience and job performance. When mechanics were questioned concerning OJT, they described it as a "pick up whatever you can" situation, which would seem to indicate that OJT is not an integrated part of maintenance shop activity.

The results also indicate that while mechanics trained in service schools perform slightly better, there is only a small difference in their performance and that of mechanics who are trained in other ways.

Also, there is only a small difference in performance between mechanics who have supervisory assignments and those who have automotive mechanic assignments. This is to be expected since a mechanic who is in a supervisory assignment is also qualified and responsible for all automotive mechanic tasks.

Finally, it is felt that the problem of developing and maintaining a highly proficient automotive mechanic is extremely complex and difficult. There are no simple answers as to exactly what makes a "proficient" mechanic. The problem as a whole is confounded with problems of personnel selection and classification, assignment procedures, the training system, and the Enlisted Evaluation System.



. 17

<sup>&</sup>lt;sup>6</sup> In a study by William C. Osborn, dealing with the performance proficiency of automotive and turret mechanics in diagnosing and repairing malfunctions.

# LITERATURE CITED AND APPENDICES

#### LITERATURE CITED

- 1. Department of the Army. Organization and Functions, Department of the Army, AR 10-5, May 1969, para 2-25(1).
- 2. Glaser, Robert, and Klaus, David J. "Proficiency Measurement: Assessing Human Performance," Psychological Principles in System Development, R.M. Gagne (ed.), Holt, Rinehart and Winston, Inc., New York, 1962.
- 3. Yellen, Ted M.I. Preliminary Validity Report of the MOS Evaluation Test for Track Vehicle Mechanic, MOS Code 632.2 (63C20), Research Study 49, U.S. Army Enlisted Evaluation Center, Fort Benjamin Harrison, Inc. 1966, p. 25.
- 4. Urry, V.W., Shirkey, E.C., and Nicewander, W.A. Preliminary Validity Report of the MOS Evaluation Test for Personnel Specialist, MOS Code 716.1, Research Study 21, U.S. Army Enlisted Evaluation Center, Fort Benjamin Harrison, Inc. 1965, p. 23.
- 5. Engel, John D. A Revised Job Requirements Inventory for General Vehicle Repairman, MOS 63C, HumRRO Research By-Product, December 1968.



#### Appendix A

# PROFICIENCY TEST FOR GENERAL VEHICLE REPAIRMAN: LIST OF EXERCISES

#### **Troubleshooting Subtest**

	Problem	Vehicle
1	Defective cooling fan clutch	M60
2	Defective wire from engine wiring junction	
	block to engine low oil pressure light	M60
3	Defective neutral switch	M113
4	Master relay will not close	M108
5	Defective fuel pump	M151A1
6	Defective hull wiring harness	M108
7	Defective wire from engine wiring junction block	
	to engine transmission oil switch	M60
8	Defective spade cylinder	M110
9	Defective clutch assembly	M35A1
10	Defective distributor capacitor	M151A1
11	Defective starter solenoid	M60
12	Defective generator regulator	M60
13	Defective starter solenoid circuit	M108
14	Defective generator to regulator wiring harness	M54A2
15	Defective instrument panel wiring harness	M60
16	Defective power plant wiring harness	M110
17	Defective generator	M151A1

#### **Corrective Action Subtest**

1	Adjust reverse and low brake bands	M60
2	Replace universal joint	M151A1
3	Adjust valve clearance	M151A1
4	Adjust reverse shift linkage	M35A1
5	Replace wheel cylinder	M151A1
6	Carburetor adjustment	M151A1
7	Repair dimmer switch wiring harness	M113
8	Replace steering gear relay lever	M54A2
9	Adjust brake bands	M60
10	Adjust clutch linkage	M151A1
11	Toe in adjustment	M151A1
12	Adjust front wheel bearing	M151A1



#### **Preventive Maintenance Subtest**

	Problem	Vehicle
1	Check Modification Work Order	M110
2	Check Modification Work Order	M110
3	Check Modification Work Order	M110
4	Check Modification Work Order	M88



#### Appendix B

# PROFICIENCY TEST FOR GENERAL VEHICLE REPAIRMAN: SAMPLE EXERCISE

#### Electrical System M54A2 (M35A1)—Five-Ton Truck

	L TOOLS AND EQUIPMENT: L.V.C.T., test adapto	or set, jumper ' 	wire 
	the subject: "The battery generator indicator show	s no charge. F	ind the
	and report it to me."		
SUBJEC	T'S PERFORMANCE:		
(1)	Started and operated the engine at 1,000-1,200 rpr	n.	<b>E</b>
(2)	Connected voltmeter from positive battery termina ground. (No increase in voltage.)	l to	
(3)	Disconnected battery connector at voltage regulator checked output from regulator. (Voltmeter from reterminal to ground) (no output)		
(4)	Disconnected generator to regulator cable at voltage connected jumper from terminal A to terminal B (and voltmeter from jumper to ground. (engine oper (no voltage)	cable side)	E
(5)	Disconnected generator to regulator cable at generator connected jumper from terminal A to terminal B or generator connector and voltmeter from jumper to	f the	TE:
(6)	(voltage now indicated)  Reported defective generator to regulator cable.		E
• •	TE: Preferred method is using adaptors and field rhe	eostat portion	
(7)		•	E
` '	Parts erroneously condemned—write in.	<del></del>	
	Did subject utilize field rheostat?	 Yes	No
	Did subject encounter any problems with		
	L.V.C.T.? (explain)	Yes	No



#### Appendix C

# PROFICIENCY TEST FOR GENERAL VEHICLE REPAIRMAN: SAMPLE SUMMARY PERFORMANCE CHECKLIST

#### SUMMARY PERFORMANCE CHECKLIST

SUBJE	CT TESTER	ITEM NO
I. SUI	BJECT DETECTED MALFUNCTION OR COMPLETED	OPERATION
	OPERLY	
CH	ECK 0. Used good procedures; knew what he was doing	
O	NE 1. Had a pretty good idea, but some guesswork	
	2. Knew very little; probably just a good guess	
	3. Knew very little; but used publication accurately	·····
		<del></del>
II. SU	BJECT FAILED	
CLI	ECK 0. Didn't know enough to get started	
-		
U	NE 1. Started but gave up very quickly	
	3. Had fairly good grasp of the problem but failed	
CHECK	ANY STATEMENTS BELOW THAT APPLY TO OR	EXPLAIN FAILURE OR
-		EATLAIN FAILURE OR
POOR	PERFORMANCE OF SUBJECT	
Procedu		
00.	Proceeding O.K. but too slow	· · · · · · · · · · · · · · · · · · ·
01.	Didn't know essential operation or step no. ()	
02.	Didn't make essential "operational" checks such as revving up en	gine,
	turning switches, etc.	
03.	Incorrectly reinstalled parts	
Location	n of Malfunction	
04.	Unable to pick out correct system to work on	
05.	Got right system but wrong components	
06.	Got right system and right components but couldn't test them .	
	pecial Tools and Test Equipment  Failed to use special equipment	
07.		
08.	Tried to use it but didn't know how to hook it up	
09.	Tried to use it but didn't know where to hook it up	
10.	Knew how and where but didn't know procedures	
11.	Used it properly but couldn't read results of checks	····· ———
<b>12</b> .	Used wrong equipment for test or check	• • • • • • • • • • • • • • • • • • •
13.	Put too much ( ) or too little ( ) stress or load on equipment	
	(Check one.)	
Use of P	Publications	
14.	Didn't use publications	
15.	Selected wrong publication	
16.	Selected right publication but couldn't find right section	· · · · · · · · · · · · · · · · · · ·
17.	Selected right publication and section but	
18.	didn't understand wiring circuit diagram	
19.	didn't understand troubleshooting diagram	
20.	didn't understand technical specifications	
III. HO	OW WAS THE SUBJECT TRAINED FOR THIS SPECIF	TC TEST PROBLEM?
0.	Has never had any training—entirely self-taught	
0. 1.		
2.	Has had one or more classes on this problem, lasting between 1/2	
4.	week. Ordnance, Tech. Rep, Battalion (or higher) NCO	'a
3.	Was taught on job by a supervisory NCO such as the Motor Serge	pant.
3. 4.	Was taught on job by another mechanic	
4. 5.	Other training (write in)	
<b>1</b>		
	(Use reverse side of this page for remarks	s.)

#### Appendix D

#### PROFICIENCY TEST FOR GENERAL VEHICLE REPAIRMAN: SAMPLE ORGANIZATIONAL MAINTENANCE PERSONNEL INVENTORY

	_	_							
Last name (PRINT)		First na	ame					ī	niti
Circle your rank.		<b>E6</b>		<b>E</b> 7		<b>E8</b>		<b>E9</b>	
Serial number			4.	Prim	ary N	ios.			
Print in the name of the	organi	ization yo	u are	now	worl	king i	n:		
What is your parent org	anizati	on, if diff	erent	from	num	ber (	5?		
Write in the number of as an automotive mecha			s of 1		ту ех	perie	ence	you have	
Write in the number of maintenance experience years	-	ave had.	s of t		vehic	ele or	ganiz	ational	
Write in the number of maintenance experience years	-	ave had:	s of v		l vehi	cle o	rgani	zational	
Make an X in the box c system on which you h	_	_	the t	уре о	f veh	icle a	nd v	ehicle	
		Take							
	Electrical	Transmission Winch, Power	Clutch	Cooling	Engine	Air-Fuel	Transfer	Differential- Front Drive	
Light Trucks							<u> </u>		
Medium Trucks									
Heavy Trucks									
Personnel Carriers									
Light Tanks									
Medium Tanks					]	]			
Tank Recovery Vehicles									



(for example, TVR Y	es No X	Where		e specifically		
where the course was tak	en.	Yes	No	Where		
Mechanic's Helper						
WVM						
TVM						
GVR						
TVR						
Advanced School Course						
If you took an advanced course, please complete the following form.						
Place		Course in wee		Subject taught		
1. ———						
2						
3				<del></del>		
3 4						
4				signment.		
4 5				signment.		
4 5 Briefly describe the kind				signment.		
4 5 Briefly describe the kind				signment.		
4 5 Briefly describe the kind				signment.		
4 5 Briefly describe the kind				signment.		

Appendix E

DIFFICULTY LEVEL (P) AND DISCRIMINATION INDEX  $(r_b)^a$ OF TEST ITEMS FOR THE

TOTAL TEST AND SUBTESTS

}	Troubleshooting			Cor	rective Ac	tion	Preventive Maintenance			
Exercise Number	P	r <sub>b</sub> for Subtest	r <sub>b</sub> for Total Test	P	r <sub>b</sub> for Subtest	r <sub>b</sub> for Total Test	P	r <sub>b</sub> for Subtest	r <sub>b</sub> for Total Test	
1	.29	.17	.19	.84	.82	.90	.65	.25	.38	
2	.34	.10	.19	.81	.68	.66	.37	.29	.40	
3	.53	.47	.49	.86	1.10	1.01	.81	.19	.36	
4	.50	.22	.40	.67	.42	.47	.68	.64	.53	
5	.71	.75	.83	.88	.60	.43	İ			
6	.79	.27	.36	.60	.19	.12				
7	.37	.11	.30	.67	.35	.57	i			
8	.37	.22	.34	1.00	0	0	l l			
9	.87	.11	.21	.83	.77	.96				
10	.55	10	.10	.83	.88	.71	<b>]</b>			
11	.27	.50	.50	.32	.27	.30				
12	.34	.48	.42	.81	.44	.57				
13	.79	.51	.64							
14	.47	.12	.15							
15	.24	.20	.24							
16	.32	.40	.33				}			
17	.45	.23	.20							

<sup>&</sup>lt;sup>a</sup>Biserial Coefficient of Correlation, corrected for spurious item-total test overlap.



#### Unclassified

Security Classification			
DOCUMENT CONTR			
(Security cleanification of title, body of abstract and indexing an			
1: ORIGINATING ACTIVITY (Corporate author)	3		URITY CLASSIFICATION
Human Resources Research Organization (Hu	<b>—</b>	Unclass	ified
300 North Washington Street	25	. GROUP	,
Alexandria, Virginia 22314			
3. REPORT TITLE			,
DEVELOPMENT OF A WORK SAMPLE CRITERION FOR	R GENERAL VEH	ICLE MEC	HANIC
4. OLECRIPTIVE NOTES (Type of report and inclusive dates) Technical Report			
5. AUTHOR(S) (First name, middle initial, lest name)			
John D. Engel			,
· · · · · · · · · · · · · · · · · · ·			
	<del></del>		
6. REPORT DATE	78. TOTAL NO. OF PAGE	ES	7b. No. OF REFS
July 1970	32		l
BE. CONTRACT OR GRANT NO.	Se. ORIGINATOR'S REPO	ORT NUMBER(5)	,
DAHC 19-70-C-0012	Technical R	eport 70	<u>-</u> 11
		cport , ;	-11
c.2Q062107A712	ST STUDE BERGET NO	'a' / Any othe	t numbers that may be seeigned
<u> </u>	this report)	(a) (Any ones	r numpere that may be evergined
d.	1		
10. DISTRIBUTION STATEMENT	<b></b>		
This document has been approved for publ:	ic release an	بهاجه ام	
its distribution is unlimited.	10 1010000	u 3a10,	
and date of the state of the st			
11. SUPPLEMENTARY HOTES	12. SPONSORING MILIT	FARY ACTIVITY	
Work Unit JOBTEST, Proficiency	Office, Chie	ef of Res	search and Development
Measurement Techniques	Department o		
•	Washington,		_
13. AGSTRACT			
<b>≜</b> *			

A work sample criterion test was developed for General Vehicle Repairman, MOS 63C3O and 63C4O. Test items covered three task categories: troubleshooting, corrective action, and preventive maintenance. Thirty-eight organizational mechanics were tested at Fort Knox, Kentucky. Data were also collected on the quality of performance, for example, use of good procedures, use of test equipment, and so forth. The study indicated that (a) the test appears to have a high degree of reliability (r=.82), (b) on the average, 60% of the test exercises were successfully completed by the 38 mechanics, (c) there was a moderate relationship between performance and length of experience, and (d) there were indications of lack of use and unfamiliarity with technical publications, and also a lack of skill in the use of special tools and equipment.

DD FORM 1473

Unclassified

Security Classification



#### Unclassified

Security Classification

4.	KEY WORDS		LINK A		LINK B		LINK C	
			ROLE	WT	ROLE	WT	ROLE	WT
Automotive Mecha	nic		Ì	]				
Criterion Development			ļ	ļ	į			
Human Performance				1	İ			
Maintenance					1	Ì		
Test Development	:							
Work Sample				Ì	ĺ			
					İ	1		
								!
	•					1		
			١					
								}
					!			1
						}	<b> </b> 	
						Ì	Ì	
		Ì						
								1
				{				}
			1					
		:						
		į			[			

Unclassified

Security Classification

#### DISTRIBUTION LIST

N LIST

I WO ARRADEEN PO ATTH TECH LIB

COMPUT USA INTELL SCH ATTH DIR DO ACADERIC OPS FT WILABIRD

COMPUT USA INTELL SCH ATTH DIR OP DOC 6 LIT FT WILABIRD

COMPUT USA CAS SCH ATTH DIR OF DOC 6 LIT FT WILABIRD

COMPUT USA CAS SCH ATTH DIR OF DOC 8 CH AMALYSIS C DOC FT GORDON

COMPUT USA CAS SCH ATTH DIR OF SCH AMALYSIS C DOC FT GORDON

COMPUT USA CAS SCH ATTH EDUC ADV FT GORDON

COMPUT USA CAS CH ATTH COT THE ACOPS GO ING DIV FT MCCEELLAN

COMPUT USA CAS CH ATTH COT THE ACOPS GO ING DIV FT MCCEELLAN

COMPUT USA SCH C THOC CITE ATTH ACOPS GO ING DIV FT MCCEELLAN

COMPUT USA FARMY CH SCH ATTH DIR OF THE STILL

COMPUT USA FARMY CH SCH ATTH DIR OF THE STILL

COMPUT USA FARMY CH SCH ATTH DIR OF THE STILL

COMPUT USA FARMY SCH ATTH DIR OF THE STILL

COMPUT USA FARMY SCH ATTH DIR OF THE STILL

COMPUT USA FARMY SCH ATTH DIR OF THE STILL

COMPUT USA FARMY SCH ATTH DIR OF THE STILL

COMPUT USA FARMY SCH ATTH DIR OF THE STILL

COMPUT USA FARMY SCH ATTH DIR OF THE STILL

COMPUT USA FARMY SCH ATTH DIR OF THE STILL

COMPUT USA FARMY SCH ATTH DIR OF THE STILL

COMPUT USA FARMY SCH ATTH DIR OF THE STILL

COMPUT USA FARMY SCH ATTH DIR OF THE STILL

COMPUT USA FARMY SCH ATTH DIR OF THE STILL

COMPUT USA FARMY SCH ATTH DIR OF THE STILL

COMPUT USA FARMY SCH ATTH DIR OF THE STILL

COMPUT USA FARMY SCH ATTH DIR OF THE STILL

COMPUT USA FARMY SCH ATTH DIR OF THE STILL

COMPUT USA FARMY SCH ATTH DIR OF THE STILL

COMPUT USA FARMY SCH ATTH DIR OF THE STILL

COMPUT USA FARMY SCH ATTH DIR OF THE STILL

COMPUT USA FARMY SCH ATTH DIR OF THE STILL

COMPUT USA CH AND SCH ATTH DIR OF THE STILL

COMPUT USA CH ATTH DIR OF THE STILL

COMPUT USA



U OF PCH LEARNING R+D CIR ATTH DIR

THUMAN SCI RES INC RELEAVA

TECH. 1900. CIR REGING DATA STAY WASHE AVW INC COLUNEUS O

CECH TO CE REGING DATA STAY WASHE AVW INC COLUNEUS O

CARTIFICATE STAY CO DITY ATTH LES STAYCH LES

GOUL C. THE COSSULTANT STATE LES STAYCH LES

GOUL C. THE COSSULTANT STATE LES STAYCH COLUNE ASS

GOUL C. THE COSSULTANT STATE LES STAYCH COLUNE ASS

GOUL C. THE COSSULTANT STAY LES STAYCH COLUNE ASS

GOUL C. THE COSSULTANT STAY LES STAYCH COLUNE ASS

HOR RES STAYCH CO DIV ATTH LES STAYCH COLUNE

HOR RESIDENCE OO DIV ATTH LES STAYCH COLUNE

HOR RESIDENCE OO THE COLUNE ASS STAYCH OF THE COLUMN

HOR RESIDENCE STAYCH COLUNE

HOR RES ARCHER CORPANY COLUNE

LOT HOR RES ON LEARNING TEACHING U D' MICH

LOT HOR RES ON LEARNING THE COLUNE

LOT HOR RES ON LEARNING THE COLUMN

CANADIAM AND STAY DEAL SHORT OF RES STAY ASSENTATION

CANADIAM AND STAY DEAL SHORT OF RES STAY ASSENTATION

CANADIAM AND STAY DEAL SHORT OF RES STAY ASSENTATION

CANADIAM AND STAY DEAL SHORT OF RES STAY ASSENTATION

CANADIAM AND STAY DEAL SHORT OF A STAY AND THE COLUMN

CARLOT OF THE FOREIGN LITASION DECR TO NOBRED MILIT ATTACHE

AREN THE LABOR OF THE STAY AND THE COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE LABOR COLUMN

CALUMN AND THE COLUMN

CALUMN AND THE COLUMN

CALUMN AND THE COLUMN

CALUMN

CALUMN AND THE COLUMN



#### **HUMAN RESOURCES RESEARCH ORGANIZATION**

300 North Washington Street . Alexandria, Virginia 22314

President Executive Vice President

Director for Business Affairs and Treasurer

Director for Operations

Director for Program Development Director for Research Design and Reporting Dr. Meredith P. Crawford

Dr. William A. McClelland

Mr. Charles W. Smith

Mr. Amold A. Heyl Dr. Robert G. Smith, Jr.

Dr. Eugene A. Cogan

#### RESEARCH DIVISIONS

HumRRO Division No. 1 (System Operations)

300 North Washington Street Alexandria, Virginia 22314

HumRRO Division No. 2

Fort Knox, Kentucky 40121 HumRRO Division No. 3

Post Office Box 5787 Presidio of Monterey, California 93940

HumRRO Division No. 4 Post Office Box 2086

Fort Benning, Georgia 31905 HumRRO Division No. 5 Post Office Box 6057 Fort Bliss, Texas 79916

HumRRO Division No. 6 (Aviation)

Post Office Box 428 Fort Rucker, Alabama 36360

HumRRO Division No. 7 (Social Science)

300 North Washington Street Alexandria, Virginia 22314

Dr. J. Daniel Lyons Director

Dr. Donald F. Haggard Director

Dr. Howard H. McFann Director

Dr. T.O. Jacobs Director

Dr. Robert D. Baldwin Director

Dr. Wallace W. Prophet Director

Dr. Arthur J. Hoehn Director

